Gambler's Ruin: Start w/ \$$i$, $\begin{cases} \text{win \$1 with prob. } p \\ \text{lose \$1 with prob. } q = 1-p \end{cases}$

Stop when hit N or go broke

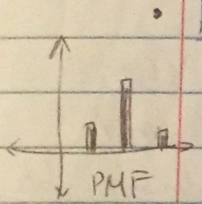
$p_i = P$ (hit N before 0, starting with i)

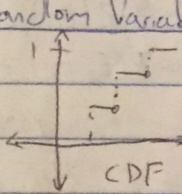
$$\Rightarrow p_i = \begin{cases} \frac{i}{N} & \text{if } p = q = \frac{1}{2} \\ \frac{1-(q/p)^i}{1-(q/p)^N} & \text{if } p \neq q \end{cases}$$

Ballot Problem: $\begin{cases} A \text{ recieves } n \text{ votes} \\ B \text{ recieves } m \text{ votes}, n > m. \end{cases}$

Chance A always ahead in vote count $= P_{n,m} = P_{n-1,m}\left(\frac{n}{n+m}\right) + P_{n,m-1}\left(\frac{m}{n+m}\right) = \frac{n-m}{n+m}$ (mbe)

- Prior: $P(A)$ = incidence
- Prior odds $\frac{P(A)}{P(A^c)}$
- Posterior: $P(A|pos)$
- Posterior odds $\frac{P(A|pos)}{P(A^c|pos)}$
- Positive Rate $P(pos|A)$
- Likelihood Ratio: $\frac{P(pos|A)}{P(pos|A^c)}$
- False Positive Rate: $P(pos|A^c)$
- Weighted Avg: $P(pos) = P(pos|A)P(A) + P(pos|A^c)P(A^c)$
- Conditional Probability is a Probability:
  - $\rightarrow P(E^c|F) = 1 - P(E|F)$
  - $\rightarrow P(E \text{ or } G|F) = P(E|F) + P(G|F) - P(EG|F)$
  - $\rightarrow P(E|G,F) = P(EG|F) / P(G|F)$
  - $\rightarrow P(EG|F) = P(G|F)\,P(E|GF)$
  - $\rightarrow 0 \leq P(E|F) \leq 1$
  - $\rightarrow P(S|F) = 1$
  - $\rightarrow E_i$'s mutually exclusive $\Rightarrow P(\bigcup_i E_i|F) = \sum_i P(E_i|F)$
- Random Variables $\rightarrow$ Probability Mass Function $p(a) = P(X=a)$
  - $\rightarrow \sum_{\text{all } x} p(x) = 1$, $0 \leq p(a) \leq 1$
  - $\rightarrow$ Cumulative Dist Function: $F(a) = P\{X \leq a\} = \sum_{\text{all } x \leq a} p(x)$
  - $\rightarrow$ Properties of CDF: $F(a_{max}) = 1$, $\lim_{a\to\infty} F(a) = 1$, $F(a) \nearrow$ stepwise





- Tail Property: $\bar{F}(a) = 1 - F(a) = P\{X > a\}$
- Mean / Expected Value: $EX = E[X] = \sum_{\text{all } x} x\,p(x)$
- 2 events E & F independent ($\perp$): $P(EF) = P(E)P(F)$, $P(E|F) = P(E)$
- Mutually exclusive - 2 outcomes cannot happen simultaneously
- Independent - outcome of 1 event has no impact on 2nd event
- $\bigcup_i F_i = F_1 \cup F_2 \cup \ldots \cup F_n$
- Assessment Biases: "Illusion of control, nonrandom "random" events, availability/recall/salience bias (recency, concreteness, egocentric attribution (funny costume) biases), anchoring bias, underestimation of situational effects.
- Misuse of Probability Rules: $P(A)$ vs $P(A^c)$ (Surgery vs. Radiation); erroneously assuming mutual exclusivity, or independence (Collins Case); ignoring independence (Monte Carlo Effect); not conditioning on important factors (Simpson's); not including all relevant info when conditioning (Monty Hall); ignoring base rate (prior), overestimation bias, assuming if posteriors for 2 groups different, priors also different (White Anglo, ~~in/out of state~~ UC)
- Misunderstanding Distributions: Ecological Fallacy (assuming mean is representative of distribution)
- ignoring pop size (Olympic medals per capita), overconfidence, good outcomes $\neq$ good decisions
- size/length bias: Random Town ex (selecting household vs resident randomly)

$X = (X_1, \dots, X_r) \sim$ Multinomial $(n, p_1, \dots, p_r)$, $\sum_i p_i = 1$.
n independent trials resulting in category i with probability p_i, r total categories.

$\Rightarrow P(X_1 = n_1, \dots, X_r = n_r)$
$= P(n_i$ results in category i, $i = 1, \dots, r)$
$= \frac{n!}{n_1! n_2! \cdots n_r!} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$, $\sum_i n_i = n$

[redacted]

IEOR 172

Notes for MT 1

- Permutations
  - # of perm's of n distinct objects $= n! = \prod_{i=1}^{n} i$
  - # of perm's of n objects, r of which are distinct
    $(n_1, n_2, \dots, n_r) = \frac{n!}{\prod_{i=1}^{r} n_i!} = \frac{n!}{n_1! n_2! \cdots n_r!}$
    (↑ alike)
- Combinations: $\binom{n}{r}$ = "n choose r" $= \frac{n!}{r!(n-r)!}$
- Binomial Theorem $(x+y)^n = \sum_{i=1}^{n} \binom{n}{i} x^i y^{n-i}$
- Multinomial # of ways to make r distinct groups of size $n_1, n_2, \dots, n_r$
  $= \frac{n!}{n_1! n_2! \cdots n_r!}$ (same as perm. #2)
- Mutually exclusive $E_i E_j = \emptyset$
- E_i's are exclusive if $\cup_i E_i = S$ ($\cup_i$ = union, S = sample space, $E \subseteq S$)
- For E_i partition (mutually exclusive & exhaustive), $P(\cup_i E_i) = 1$
- E "$\subseteq$" F ("subset of") $\Rightarrow P(E) \leq P(F) \Rightarrow P(EF) \leq P(F)$
- $P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3) - P(E_1E_2) - P(E_1E_3) - P(E_2E_3) + P(E_1E_2E_3)$

Rules
- $\ominus$ $P(E^c) = 1 - P(E)$
- $\oplus$ $P(E \cup F) = P(E) + P(F) - P(EF)$
  $EF = \emptyset \Rightarrow P(E \cup F) = P(E) + P(F)$
- $\div$ $P(E|F) = P(EF)/P(F) = P(F|E)P(E)/P(F)$
- $\otimes$ $P(EF) = P(F)P(E|F)$
  $P(E_1 E_2 \dots E_n) = P(E_1) P(E_2|E_1) P(E_3|E_1E_2) \dots$
- TP for F_i a partition $P(E) = \sum_i P(F_i) P(E|F_i)$ [ie. $= P(F)P(E|F) + P(F^c)P(E|F^c)$]
- Bayes's for F_j a partition $P(F_j|E) = \frac{P(F_j)P(E|F_j)}{\sum_i P(F_i)P(E|F_i)}$ ← TP

• Simpson's paradox: M+F admission stats ex.

- For A = event & pos = test positive (evidence in favor of A), assuming $P(pos|A) \geq P(pos|A^c)$ (true if false pos + false neg both $\leq 50\%$):
  - → $P(pos|A) \geq P(pos) \geq P(pos|A^c)$
  - → $P(A) \approx 0 \Rightarrow P(pos) \approx P(pos|A^c)$
  - → $P(A|pos) \leq P(A)/P(pos|A^c)$
  - → $P(A) \approx 0$, $P(pos|A) \approx 1 \Rightarrow P(A|pos) \approx P(A)/P(pos|A^c)$
  - → $P(pos|A^c) \gg P(A) \Rightarrow P(pos) \gg P(A)$ [estimation bias]
  - → $P(pos|A^c) \geq P(A) \approx 0 \Rightarrow P(pos\ A^c) \geq P(pos\ A)$
  - → $\frac{P(A|pos)}{P(A^c|pos)} = \frac{P(A)\ P(pos|A)}{P(A^c)\ P(pos|A^c)}$

$\ldots V(X) + Var(Y) + 2\,Cov(X, Y)$

* Properties of Binomial: • Mode = floor of $(n+1)p$.
• $X \sim Bin(n,p) \doteq Normal(\mu, \sigma^2) = Normal(np, np(1-p))$
* for n large

$X \sim Zeta(\alpha)$, $\alpha > 0$
$P_X(k) = \frac{c}{k^{\alpha+1}}$ for $k=1,2,\ldots$
where c is such that the probability sums to 1.

[redacted]

IEOR 172

Notes for MT 2

*mean = balance density : $E[X-\mu] = 0$

Law of large numbers : if x observed a large # of times, avg. of observed values go to EX.

Discrete PMF : $p(a) = P\{X = a\}$ $\qquad E[X] = \sum_x x\,p(x)$
CDF : $F(a) = P\{X \leq a\} = \sum_{x \leq a} p(x)$ $\qquad E[g(x)] = \sum_x g(x)p(x)$

Continuous PDF : $f(x)$; $P\{X = x\} = 0$ prob. = area under $f(x)$ $\qquad E[X] = \int_{-\infty}^{\infty} x f(x)\,dx$
CDF : $\int_{-\infty}^{\infty} F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)$ $\qquad E[g(x)] = \int_{-\infty}^{\infty} g(x) f(x)\,dx$

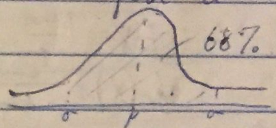
Newsvendor ex profit \$B, loss for unsold \$L. $*s$ = smallest s such that $F(s) \geq \frac{B}{B+L}$
s^* satisfies $F(s^*) = \frac{B}{B+L}$ (s^* = optimal stocking level)

For all RVs: $E(\sum_i a_i X_i + b) = \sum_i a_i EX_i + b \Rightarrow E(X+Y) = EX + EY$
$Var(X) = E[(X-EX)^2] = E(X^2) - E[X]^2$
$Var(aX+b) = a^2 Var(X)$
If $X \perp Y \Rightarrow Var(X+Y) = Var(X) + Var(Y)$, if X_i's $\perp \Rightarrow Var(\sum_i a_i X_i + b) = \sum_i a_i^2 Var(X_i)$
$SD(X) = \sqrt{Var(X)}$ $\qquad SD(aX) = a\,SD(X)$

Jensen's Inequality $g(x)$ convex $(g''(x) > 0) \Rightarrow E[g(X)] \geq g[E(X)]$

Chebyshev's Inequality: For any R.V. X with mean μ, variance σ^2, for any k
$P\{|X-\mu| \geq k\} = \frac{\sigma^2}{k}$ $\Rightarrow$ For any distribution...
• ~~at least 75% of prob within~~ 2 SD of mean
• at least 89% of prob. within 3 SD of mean

Empirical Rule: For bell-shaped distribution ~68% of prob within 1 SD of mean



~95% " 2 "
~99.7% " 3 "

CLT: suppose X_i's iid (indep & identically distributed) with mean μ & SD σ.
then sample mean $\bar{X} = \sum_{i=1}^{n} X_i/n \approx N(\mu, \frac{\sigma^2}{n})$. <bigger $n \rightarrow$ more like a bell>

Bernoulli: $X \sim Bernoulli(p)$ X is indicator of success in 1 trial.
$X = \begin{cases} 1 & wp\ p \\ 0 & wp\ 1-p \end{cases}$ $\qquad EX = p$ $\qquad Var(X) = p(1-p)$

* **Binomial:** $X \sim Bin(n,p)$ # of successes in n indep. trials, each trial with success prob p.
$X = \sum_i Y_i \quad Y_i \sim Bern(p)$ $\qquad EX = np$
$p(i) = P\{X = i\} = \binom{n}{i} p^i (1-p)^{n-i}$ $\quad i = 0, \ldots, n$ $\qquad Var(X) = np(1-p)$
$X_1 \sim Bin(n_1, p)$, $X_2 \sim Bin(n_2, p) \Rightarrow X_1 + X_2 \sim Bin(n_1 + n_2, p)$ <X_1, X_2 indep.>
$X \sim Bin(n,p)$ n large, $p \approx \frac{1}{2} \Rightarrow X \doteq N(np, np(1-p))$ <need larger n for p further from $\frac{1}{2}$>

Poisson: $X \sim Bin(n,p)$ n large, p small $\Rightarrow X \sim Poisson(\lambda)$ $[\lambda = np]$
Poisson better approx. than normal approx if p small.
if p close to 1 $\Rightarrow (1-p)$ small $\Rightarrow$ approx. failures with Poisson(λ) with $\lambda = n(1-p)$.
$X_i \sim Poisson(\lambda)$ / $p(i) = e^{-\lambda} \frac{\lambda^i}{i!}$ $i = 0, 1, 2$ / $EX = \lambda = Var(X)$

Thm: $X_i \sim exp(\lambda_i)$, $\perp$ [redacted]

$\Rightarrow \min(X_i) \sim exp(\sum \lambda_i)$

Poisson Paradigm: $X \sim Poisson(\lambda)$ a good approx. for counting # of success out of n (for n large, $\lambda = np$)
$X = \sum I_i$ when (Binomial) (sum of Bernoulli)
→ there's weak dependency in trials (so I_i's weakly dependent) AND
→ the p's aren't the same (so $I_i \sim Bern(p_i)$); ok as long as all p_i's are small.

ex) # winners in lottery, # overtime games in NFL, # people entering store 9–10 on weekdays, etc...

Poisson Process: $\{N(t)\}$ where $N(t)$ = # arrivals in time t, λ = arrival rate, $N(t) \sim Poisson(\lambda t)$

time between arrivals $\sim exp(\lambda)$

$\Rightarrow \{N(t)\}$ a poisson process if: (1) $P\{N(h)=1\} = \lambda h + o(h)$
(2) $P\{N(h) \geq 2\} = o(h)$
} for h small $N(h) \approx Bern(\lambda h)$
(3) # of events/arrivals in non-overlapping intervals indep.

Poisson Splitting: $N \sim Poisson(\lambda)$, N's split into 2 groups, with prob p it will go to group 1.
X = group 1, Y = group 2 $\Rightarrow$ $X \sim Poisson(\lambda p)$ & $Y \sim Poisson(\lambda(1-p))$ where $X \perp Y$

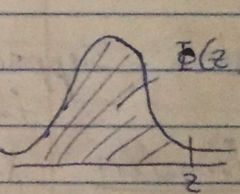
Geometric: $X \sim Geom(p)$ # of trials until 1st success in indep. trials with success prob. p
$p(n) = (1-p)^{n-1} p$ $\quad EX = \frac{1}{p}$
$F(n) = 1 - (1-p)^n$ $\quad Var(X) = \frac{1-p}{p^2}$

Negative Binomial: $X \sim NegBin(r, p)$ # of trials to get r successes where trials indep. each with success prob. p.
$P\{X=n\} = P\{r^{th}$ success on n^{th} trial$\}$ $= \binom{n-1}{r-1} p^{r} (1-p)^{n-r}$ $\quad EX = \frac{r}{p}$ $\quad Var(X) = \frac{r(1-p)}{p^2}$

Hypergeometric: $P\{X=i\} = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}$
$X \sim Hypergeom(n, m, N)$ # of white balls in random selection from urn where n balls selected & there are m white balls & N total balls. RV's
$EX = \frac{mn}{N}$ $\quad Var(X) = np(1-p)\left[1 - \frac{n-1}{N-1}\right]$ where $p = \frac{m}{N}$ <w/o replacement>

Uniform: $X \sim Unif(a,b)$ $\quad f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a<x<b \\ 0 & \text{otherwise} \end{cases}$ $\quad EX = \frac{b-a}{2}$ $\quad Var(X) = \frac{(b-a)^2}{12}$

Normal: $X \sim Norm(\mu, \sigma^2)$ $\quad f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}$
$EX = \mu$, $Var = \sigma^2$



Standard Normal: $Z = \frac{X-\mu}{\sigma}$, $Z \sim N(0,1)$. $\quad F(z) = \Phi(z) = P(Z \leq z)$

Exponential (*memoryless*): $X \sim exp(\lambda)$ $\quad f(x) = \lambda e^{-\lambda x}$, $x \geq 0$ $\quad EX = \frac{1}{\lambda} = SD(X)$
$\bar{F}(x) = e^{-\lambda x}$ $\quad Var(X) = \frac{1}{\lambda^2}$

Memoryless: eg. $P(X \geq t+x \mid X \geq t) = P(X \geq x)$ $\quad$ t t+x

Joint CDF *: $F_{X,Y}(x,y) = P\{X \leq x, Y \leq y\}$ $\quad f_X(x) = \int_a^b f(x,y)\,dy$. $\quad \alpha^{th}$ percentile $= P\{Z \leq a\} = \alpha$

Marginal CDF: $F_X(x) = P\{X \leq x\} = P\{X \leq x, Y \leq \infty\} = \lim_{y \to \infty} F_{X,Y}(x,y)$

Conditional: PMF: $p_{X|Y}(x|y) = P\{X=x \mid Y=y\} = \frac{p_{X,Y}(x,y)}{p_Y(y)}$ (Disc.) $\Rightarrow \left(\frac{\text{joint}}{\text{marginal}}\right)$
{Cont. PDF: $f_{X|Y} = \frac{f_{X,Y}(x,y)}{f_Y(y)}$}

Independence: $F_{X,Y}(x,y) = F_X(x) F_Y(y)$ $\forall x,y$
(Disc.) $p_{X,Y}(x|y) = p_X(x) p_Y(y)$ $\forall x,y$ $\iff p_{X|Y}(x,y) = p_X(x)$ $\forall y$ & vice versa for $p_{Y|X}$.

To show X, Y not independent, show $p_{X,Y}(a,b) \neq p_X(a) p_Y(b)$ OR show conditional $\neq$ marginal

$P\{X=x, Y=y\}$:

X \ Y	0	1	2
0	$p_{X,Y}(0,0)$	...	
1	...		
2			

Find $p_X(x)$ by summing over rows
" $p_Y(y)$ " columns

* {joint PMF: $F_{X,Y} = \sum_{a \leq x} \sum_{b \leq y} p_{X,Y}(a,b)$}
{marginal PMF: $p_X(x) = P\{X=x\} = \sum_y p(x,y)$}

- $P\{a_1 < X < a_2, b_1 < Y < b_2\} = F_{X,Y}(a_2,b_2) - F_{X,Y}(a_1,b_2) - F_{X,Y}(a_2,b_1) + F_{X,Y}(a_1,b_1)$

[redacted]

- $Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$
- $X \perp Y \Rightarrow Cov(X,Y) = 0$

Thm: $X_i \sim exp(\lambda_i)$, $\perp$ $\Rightarrow \min(X_i) \sim exp(\sum \lambda_i)$

IEOR 172

Notes for Final

$E[\sum^{N} X_i] = ENE[X_i]$

Sums of RV:
- $Z = X + Y$: $Z = z, X = x \iff X = x, Y = z - x$
- $p_Z(z) = \sum_x p_{X,Y}(x, z-x)$ <discrete>
- $f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x)\,dx$. <continuous>

Special Cases:
- $X \sim Poisson(\lambda_1)$, $Y \sim Poisson(\lambda_2)$, $X \perp Y \Rightarrow X + Y \sim Poisson(\lambda_1 + \lambda_2)$.
- $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, $X \perp Y \Rightarrow X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
- By CLT: addition/~~[illegible]~~ average of non-bell shaped distributions = approx. bell shaped for a large enough

Expected Values:
- Discrete: $E[g(X,Y)] = \sum_x \sum_y g(x,y)\,p_{X,Y}(x,y)$
- Continuous: $E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy$
- $X \perp Y \Rightarrow E[g(X)h(Y)] = E[g(X)]\,E[h(Y)]$

Variance:
- $Var(g(X,Y)) = E[\{g(X,Y) - E[g(X,Y)]\}^2] = E[g(X,Y)^2] - (E[g(X,Y)])^2$

Covariance:
- $Cov(X,Y) = E[(X - EX)(Y - EY)] = E(XY) - E(X)E(Y)$

Properties of Covariance:
- $Cov(X,Y) = Cov(Y,X)$
- $Cov(X, Y+Z) = Cov(X,Y) + Cov(X,Z)$
- $Cov(X,X) = Var(X)$
- $Cov(\sum_i X_i, \sum_j Y_j) = \sum_i \sum_j Cov(X_i, Y_j)$
- $Cov(aX, Y) = a\,Cov(X,Y) \Rightarrow Cov(aX, aY) = a^2 Cov(X,Y)$
- $Cov(X,Y)$: >0 (positive correlation); <0 (negative correlation); $=0$ (no correlation)

Correlation:
- Correlation measures linear relationship (+: tend to be large together, −: when one large, other tends to be small, no corr: no linear relationship between RVs)
- $\rho(X,Y) = \frac{Cov(X,Y)}{SD(X)\,SD(Y)}$, $-1 \le \rho(X,Y) \le 1$.
- $\rho(X,Y)$: 1 (perfect + corr. <eg. $Y = 2X$>), −1 (perfect − corr. <eg. $Y = -2X$>), 0 ($Cov(X,Y) = 0 \iff$ uncorrelated)
- independence $\Rightarrow$ uncorrelated (but not reverse!)

Formula:
- $Var(\sum_i a_i X_i + b) = \sum_i a_i^2 Var(X_i) + \underbrace{\sum_{i,j,\,i\neq j} a_i a_j Cov(X_i, X_j)}_{=0 \text{ if } X_i\text{'s pairwise indep. (ie. } X_i \perp X_j,\ \forall i \neq j)}$
  $= \sum_{i,j} a_i a_j Cov(X_i, X_j)$
- If a_i & $Var(X_i)$ are the same for all i & $Cov(X_i, X_j)$ is the same for all i & j and there are n X_i's, $\Rightarrow Var(\sum a X_i + b) = n a^2 Var(X_i) + n(n-1) a^2 Cov(X_i, X_j)$

Conditional Expectation:
- continuous: $E[X|Y=y] = \int_{-\infty}^{\infty} x\, f_{X|Y}(x|y)\,dx$, $E[g(X)|Y=y] = \int_{-\infty}^{\infty} g(x)\, f_{X|Y}(x|y)\,dx$
- discrete: $E[X|Y=y] = \sum_x x\, p_{X|Y}(x|y)$, $E[g(X)|Y=y] = \sum_x g(x)\, p_{X|Y}(x|y)$
- $X \perp Y \Rightarrow E[X|Y=y] = EX$ for all y.

Computing:
- $Z = E[X|Y]$ a RV. $E[X] = E[E[X|Y]] = \begin{cases} \int_{-\infty}^{\infty} E[X|Y=y] f_Y(y)\,dy & \text{<continuous>} \\ \sum_y E[X|Y=y]\, p_Y(y) & \text{<discrete>} \end{cases}$

Gambler's Ruin:
- expected time to ruin: $ET_i = E[\text{Time to ruin} \mid \text{start with } i] = \begin{cases} \infty & \text{if } p \ge \frac{1}{2} \\ \frac{i}{1-2p} & \text{if } p < \frac{1}{2} \end{cases}$

... Sample MT #1. (for MT 1)

1. a. $8!$ ways

$Var(X) = E[Var(X|N)] + Var(E[X|N])$

for $\sum_{i=k+1}^{n-1} \frac{1}{i} > 1$, keep going

Stop when $\sum_{i=k+1}^{n-1} \frac{1}{i} \leq 1$.

k^* is first k where $\sum_{i=k+1}^{n-1} \leq 1$.

Prop 3: $Var(X_i) = P\{X_i = 1\}(1 - P\{X_i = 1\})$

$Var(N) = E[N^2] - (E[N])^2$

$Cov(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j]$

$E[X_i X_j] = P\{X_i X_j = 1\}$

$= P\{X_i = 1\} P\{X_j = 1 | X_i = 1\}$

* from extra final practice Q's

Random sum of RV: • X_i's have mean $EX_i = \mu$ (not necessarily iid). N integer valued, $N \perp X_i$, $Y = \sum_{i=1}^{N} X_i$.

$\Rightarrow E[Y] = E[\sum_{i=1}^{N} X_i] = E[E(\sum_{i=1}^{N} X_i | N)] = E[N E(X_i)] = E[N\mu] = \mu E[N]$

Computing Probabilities: Let $X = I\{A \text{ occurs}\} \Rightarrow EX = P(X=1) = P(A)$, so

$$P(A) = E[X] = E[E[X|Y]] = \begin{cases} \int_{-\infty}^{\infty} P(A|Y=y) f_Y(y)\,dy & \text{<continuous>} \\ \sum_y P(A|Y=y) P(Y=y) & \text{<discrete>} \end{cases}$$

Best Prize: • See n prizes one at a time, want to maximize P(chosen candidate is best). Optimal policy is to let k^* go by and then select next prize that is best of those seen so far.

- $P(\text{getting best} \mid \text{let } k \text{ go by}) = P_k(B) = \frac{k}{n}\sum_{i=k}^{n-1}\frac{1}{i}$, $P_0(B) = \frac{1}{n}$. $P_{k^*} = \frac{k^*}{n}\sum_{i=k^*}^{n-1}\frac{1}{i}$
- To find optimal k^*, start with $k=0$, compute $P_k(B)$, $\uparrow k$ until $P_k(B) \downarrow$
- If n large, $k^* \approx \frac{n}{e} = 0.37n$. $P_{k^*}(B) \approx \frac{1}{e} \approx 0.37$

Cond. Var. Formula: $Var(X) = E[Var(X|Y)] + Var[E(X|Y)]$

Japan (cater) ◦ Prop 1: $X \perp Y \Rightarrow Var(XY) = E(Y^2)Var(X) + E(X)^2 Var(Y) = E(X^2)Var(Y) + E(Y)^2 Var(X)$

Restaurant (French) • Prop 2: X_i's uncorrelated with same mean & variance, $N \perp X_i$'s for all i, N = nonnegative integer RV

$\Rightarrow Var(\sum_{i=1}^{N} X_i) = EN\,Var(X_i) + (EX_i)^2 Var(N)$

(w/o replace) U.S. ◦ Prop 3: $N \perp X_i$ for all i, X_i's not independent (not uncorrelated), N a nonnegative integer RV, EX_i and $Var(X_i)$ same for all i, $Cov(X_i, X_j)$ same for all i, j

$\Rightarrow Var(\sum_{i=1}^{N} X_i) = EN\,Var(X_i) + E(X_i)^2 Var(N) + (E(N^2) - E(N))\,Cov(X_i, X_j)$

Prediction: • Predict a RV Y with a single # c. Optimal guess c^* depends on objective.

- Prop 1: best guess c^* for minimizing mean squared error (MSE) $h(c) = E[(Y-c)^2]$ is $c^* = EY$.

  Cor: At c^*, MSE is $h(c^*) = Var(Y)$. $E[(Y|X)]$

- Prop 2: best guess c^* for minimizing mean absolute error (MAD) $h_A(c) = E[|Y-c|]$ is the median, ie. c^* such that $F(c) = 0.5$. MSE more sensitive to extreme values. Guessing median (minimizing MAD) considered more representative.

- Prop 3: c^* such that $F(c^*) = \frac{b}{b+h}$ minimizes newsvendor cost $h_V(c) = E[h(c-Y)^+ + b(Y-c)^+]$ (ie. penalty for overestimating different from penalty for underestimating).

- Prop 4: For same penalty cost for any wrong guess, ie. minimizing $P\{\text{wrong}\} = P\{c \neq Y\}$, c^* = mode.

- Now X, Y related, & we know X. We'll predict Y with $g(X)$.

- Prop 5: Best guess $g^*(x)$ for minimizing $h(g(X))$ (MSE given X) is $g^*(x) = E[Y|X]$

  Cor: At $g^*(x)$, MSE given X is $h(g^*(x)) = E[Var(Y|X)]$.

for nonnegative RV Y (integer), $Var(YX) = Var(\sum_{i=1}^{Y} X) = E[Y^2]Var(X) + (EX)^2 Var(Y)$

* Get better estimate for Y if we use X: $Var(Y) = E[Var(Y|X)] + \underbrace{Var[E(Y|X)]}_{\geq 0} \geq E[Var(Y|X)]$

Packet Transm. • $P\{\text{loss during busy period}\} = P\{L\}$: condition on X = # arrivals at start of B.P. ($\neq 0$)

- $P\{X=0\} = 0$, $P\{X=1\} = P\{A_i = 1 | A_i > 0\} = \frac{p_1}{p_1+p_2+p_3} = \frac{p_1}{1-p_0}$

- $P\{\text{losses}\} = P\{L|X=1\}P\{X=1\} + \cdots + P\{L|X\geq 3\}P\{X \geq 3\} = P\{L_1\}\frac{p_1}{1-p_0} + P\{L_2\}\frac{p_2}{1-p_0} + 1\cdot\frac{p_3}{1-p_0}$

  $= 0\cdot p_0 + P\{L_1\}p_1 + P\{L_2\}p_2 + 1\cdot p_3$